THE BANDWIDTH OF CATERPILLAR GRAPHS

Zevi Miller
Department of Mathematics and Statistics
Miami University, Oxford, Ohio 45056

1. Introduction

Let G be a graph with no loops or multiple edges. We let V(G) and E(G) be the point set and edge set of G respectively, and we suppose G has n points. A <u>labelling</u> ℓ of G is an injection ℓ: V(G) $\to$ $\mathbb{Z}^+$ of V(G) into the positive integers. If v ε V(G), we call $\ell(v)$ the <u>label of v</u>. Unless otherwise indicated explicitly or by context the range of ℓ (that is, the set $\{\ell(v): v \; \varepsilon \; G\}$) will be the set $\{1,2,\cdots,n\}$. We define the <u>absolute value</u>, $|\ell|$, of ℓ by $|\ell| = \max\{|\ell(v) - \ell(w)|: vw \; \varepsilon \; E(G)\}$. The <u>bandwidth</u>, <u>B(G)</u>, of G is given by $B(G) = \min\{|\ell|: \ell$ a labelling of G}.

The problem of determining the bandwidth of a graph is related to problems of matrix computations confronted in engineering. These problems often involve the inversion of a given 0 - 1 matrix M by the use of certain algorithms which work most efficiently when all the ones lie in a narrow band about the main diagonal. If M is symmetric with diagonal having all zeros then it may be viewed as the adjacency matrix of a graph. The problem of finding a matrix equivalent to M (by simultaneous row and column interchanges) with ones placed in as narrow a band as possible about the diagonal is then equivalent to the problem of determining the bandwidth of the graph which M represents.

Since mention of this problem in graph theoretical terms in [7] there has developed a growing literature on bandwidth devoted to exact results, efficient algorithms, and NP-completeness results [4, 5, 8 to name a few]. A summary of known results and bibliography to the present may be found in [3]. It is the object of this paper to develop a O(kn) algorithm for the determination of B(C), where C is an arbitrary caterpillar graph on n points having k nonendpoints.

CONGRESSUS NUMERANTIUM, Vol. 33 (1981), pp. 235-252.

We now introduce the notation to be used in this paper. Let P_k be the path graph on the k points $v_1, v_2, \cdots, v_k$ as traversed in succession from the endpoint v_1 to the endpoint v_k. Now let m_i, $1 \leq i \leq k$, be a sequence of nonnegative integers with $m_1 > 0$ and $m_k > 0$. We then define the *caterpillar* $C = \underline{C(m_1, m_2, \cdots, m_k)}$ to be the tree obtained from P_k by adding $\sum_{i=1}^{k} m_i$ points and joining m_i of them to v_i with edges, for each i, $i \leq 1 \leq k$. The caterpillar C(2,5,2) is illustrated in Figure 1. The points $v_1, v_2, \cdots, v_k$ of C are then the only points in C having degree larger than 1, and they will be called the *links* of C. We let the *endpoint degree* of a link in C be the number of endpoints adjacent to that link. Now let G be a graph and let $S \subseteq V(G)$. We then define $\langle S\rangle$, the *graph induced by S*, by $V(\langle S\rangle) = S$ and $E(\langle S\rangle) = \{vw: v \in S, w \in S\}$. Notice that the caterpillar $C(m_1, m_2, \cdots, m_k)$ has subcaterpillars $\underline{C_i}$ for each i, $1 \leq i \leq k$, defined by $C_i = \langle v_j, \{\text{endpoints of } C \text{ adjacent to } v_j\}, 1 \leq j \leq i\rangle$. Thus we have $C_i \cong C(m_1, m_2, \cdots, m_i)$. We then define $\underline{\tilde{C}_{i-1}}$ for $i \geq 2$ as $\tilde{C}_{i-1} = \langle V(C_{i-1}) \cup \{v_i\}\rangle$, so that $\tilde{C}_{i-1} \cong C(m_1, m_2, \cdots, m_{i-2}, m_{i-1} + 1)$. Also we let $S_i = |C_i| = \sum_{j=1}^{i} (1 + m_j)$.

Suppose ℓ is a labelling of the graph G. If H is a subgraph of G we denote by $\ell|H$ the labelling of H obtained by restricting ℓ to H, so that $|\ell|H|$ would be the absolute value of $\ell|H$. We call ℓ a *monotone labelling* (or *m-labelling* for short) of the caterpillar $C = C(m_1, m_2, \cdots, m_k)$ if for each i, $1 \leq i \leq k$, v_i and its adjacent endpoints receive under ℓ the labels $S_{i-1} + 1, S_{i-1} + 2, \cdots, S_i$ in some order. An m-labelling of C(2,5,2) is illustrated in Figure 1a. Any m-labelling ℓ of the caterpillar C is uniquely determined (up to automorphisms of C and transformations $\ell(x) \to |C| - \ell(x) + 1$ for all $x \in V(C)$) by its values at the links of C since any two endpoints adjacent to the same link of C are similar under the automorphism group of C. We call the values $\ell(v_1), \ell(v_2), \cdots, \ell(v_k)$ the *pointers* of the m-labelling ℓ. Now if $a_1, a_2, \cdots, a_k$ is a sequence of positive integers satisfying $S_{i-1} + 1 \leq a_i \leq S_i$, where $S_0 = 0$, then we refer to the m-labelling of C having pointers $\ell(v_1) = a_1$, $\ell(v_2) = a_2, \cdots$, $\ell(v_k) = a_k$ as the m-labelling of C defined by the pointers $a_1, a_2, \cdots, a_k$. Although the k link points of C are denoted by $v_1, v_2, \cdots, v_k$, we will on occasion refer to the endpoint adjacent to v_1 with label 1 by v_0 and to the endpoint adjacent to v_k with label $|C|$ by v_{k+1}, if such a pair of endpoints exist.

For any graph theoretic terminology not defined in this paper we refer the reader to any of the texts [1,2,6].

2. The Algorithm

We begin with the following algorithm for obtaining a monotone labelling L of the arbitrary caterpillar $C = C(m_1, m_2, \cdots, m_k)$. We get L by successively constructing monotone labellings L_i of the subcaterpillars $C_i = C(m_1, m_2, \cdots, m_i)$ $i \leq i \leq k$. Specifically, we will recursively construct a set of integers $g(i,j)$, $1 \leq i, j \leq k$, and the reader may think of L_i as the monotone labelling of C_i induced by letting $L_i(v_j) = g(i,j)$ for $1 \leq j \leq i$. In the process we construct the integers f_i, $1 \leq i \leq k$, satisfying $f_i = |L_i|$. It will eventually be shown that $|L_i| = B(C_i)$, so that in particular $|L_k| = |L| = B(C)$ and we will have found a polynomial algorithm for obtaining the bandwidth of a caterpillar.

Algorithm CATBAND

Initialize by setting $\lambda_0 = f_1 = \{\frac{m_1}{2}\}$ and $g_0(1,1) = g(1,1) = 1 + \{\frac{m_1}{2}\}$.

Now let i satisfy $1 < i \leq k$, and suppose we have defined integers f_j and $g(i-1,j)$, $1 \leq j \leq i-1$, with the property $g(i-1,i-1) + f_{i-1} \geq S_{i-1}$. We then define the integer f_i and the set of integers $g(i,j)$, $1 \leq j \leq i$, as follows.

Our first task is to reduce to the case $g(i-1,i-1) + f_{i-1} > S_{i-1}$.

This will be done in step 1 by adjusting the values of f_{i-1} and $g(i - 1,j)$, $1 \leq j \leq i - 1$, to new values λ_{i-1} and $h(i - 1,j)$ respectively which satisfy $h(i - 1,i - 1) + \lambda_{i-1} > S_{i-1}$.

1. Suppose $g(i - 1,i - 1) + f_{i-1} = S_{i-1}$. We set $h(i - 1,1) = \min\{g(i - 1,1) + 1, S_1\}$ and recursively for $2 \leq j \leq i - 1$ $h(i - 1,j) = \min\{h(i - 1,j - 1) + 1 + f_{i-1}, S_j\}$. Also let $\lambda_{i-1} = f_{i-1} + 1$. Go to step 3.

2. If $g(i - 1,i - 1) + f_{i-1} > S_{i-1}$, let $\lambda_{i-1} = f_{i-1}$ and set $h(i - 1,j) = g(i - 1,j)$, $1 \leq j \leq i - 1$.

Note that at the end of step 2, we have obtained integers λ_{i-1} and $h(i - 1,j)$, $1 \leq j \leq i - 1$, satisfying $h(i - 1,i - 1) + \lambda_{i-1} > S_{i-1}$. These integers represent an adjusted labelling L'_{i-1} of C_{i-1}, namely, the m-labelling of C_{i-1} with pointers $L'_{i-1}(v_j) = h(i - 1,j)$, $1 \leq j \leq i - 1$, which satisfies $|L'_{i-1}| = \lambda_{i-1}$. Under the inductive assumption that $|f_{i-1}| \leq |\ell|$ for any m-labelling ℓ of C_{i-1}, it follows that λ_{i-1} is the smallest possible value of $L|\tilde{C}_{i-1}$ for any m-labelling L of C_i. Our object will now be to obtain a m-labelling L_i of C_i whose absolute value is as little above λ_{i-1} as possible. In case $S_i - (h(i - 1,i - 1) + \lambda_{i-1}) < \lambda_{i-1}$ we can actually find an L_i with $|L_i| = \lambda_{i-1}$, and step 3 does this. If $S_i - (h(i - 1,i - 1) + \lambda_{i-1}) \geq \lambda_{i-1}$, then step 4 constructs m-labellings ℓ_t of C_i for each integer $t \geq 0$ satisfying $|\ell_t|\tilde{C}_{i-1}| = \lambda_{i-1} + t$. The crux of the algorithm is to find a "compromise" between m-labellings ℓ of C_i having a low value of $|\ell|\tilde{C}_{i-1}|$ and a high $|S_i - \ell(v_i)|$ (as exemplified by the ℓ_t with low t), and those with a high value of $|\ell|\tilde{C}_{i-1}|$ and low $|S_i - \ell(v_i)|$ (as in the ℓ_t with high t). the compromise reached is the m-labelling L_i of C_i with pointers $L_i(v_j) = g(i,j)$, $1 \leq j \leq i$, where the $g(i,j)$ are constructed in step 5.

3. If $S_i - (h(i-1,i-1) + \lambda_{i-1}) \le \lambda_{i-1}$, then let $g_0(i,i) = \min\{h(i-1,i-1) + \lambda_{i-1}, S_i\}$ and $g_0(i,j) = h(i-1,j)$ for $1 \le j \le i-1$. Go to step 5.

4. If $S_i - (h(i-1,i-1) + \lambda_{i-1}) > \lambda_{i-1}$, then let $g_0(i,i) = h(i-1,i-1) + \lambda_{i-1}$ and $g_0(i,j) = h(i-1,j)$, $1 \le j \le i-1$, and define integers $g_t(i,j)$, $t \ge 1$, as follows. Let $g_t(i,1) = \min\{g_{t-1}(i,1)+1, S_1$ and for $1 < j \le i$ we let $g_t(i,j) = \min\{g_t(i,j-1) + \lambda_{i-1} + t, S_j\}$.

5. Now let $s_i \ge 0$ be the first integer for which $S_i - g_{s_i}(i,i) \le \lambda_{i-1} + s_i$. Then define $g(i,j) = g_{s_i}(i,j)$ for $1 \le j \le i$, and let $f_i = \lambda_{i-1} + s_i$.

Now repeat the algorithm with i and $i-1$ replaced by $i+1$ and i respectively. The algorithm ends after it has been carried out for $i = k$. (It will turn out that $f_k = B(C)$.)

We illustrate how CATBAND operates in finding an m-labelling of $C(2,5,2)$. For each i, $1 \le i \le 3$, we list the sequence $(g_t(i,1), g_t(i,2), \cdots, g_t(i,j))$, $1 \le j \le i$, for each t, $0 \le t \le s_i$, as computed by CATBAND.

(2) - Initialization

(3,5) - steps 1 and 4 of first iteration

(3,6) - steps 4 and 5 of first iteration

(3,7) - step 1 of second iteration

(3,7,11) - steps 3 and 5 of second iteration

Observe that we have $s_2 = 1$ and $s_3 = 0$.

The essence of the algorithm, as embodied in steps 4 and 5, is to construct an m-labelling ℓ_t of C_i with pointers $\ell_t(v_j) = g_t(i,j)$, $i \le j \le i$, for each t satisfying $0 \le t \le s_i$, and such that $|\ell_t|\tilde{C}_{i-1}| = \lambda_{i-1} + t$. We also want each of the pointers of ℓ_t to be as large as possible (subject to the constraint on $|\ell_t|\tilde{C}_{i-1}|$) so as to minimize $S_i - \ell_t(v_i)$. Note that as long as $S_i - \ell_t(v_i) > |\ell_t|\tilde{C}_{i-1}|$ we have $|\ell_t| \ge S_i - \ell_t(v_i) > |\ell_t|\tilde{C}_{i-1}|$.

In this case we may as well increase the pointers $\ell_t(v_j)$, $1 \le j \le i$, (thereby obtaining the new labelling ℓ_{t+1}) allowing if necessary an increase by 1 in the distance between any two successive ones. This increase in successive differences of course makes $|\ell_{t+1}|\tilde{C}_{i-1}|$ larger by 1 than $|\ell_t|\tilde{C}_{i-1}|$, but since $|\ell_t|\tilde{C}_{i-1}| + 1 \le |\ell_t|$ the value of $|\ell_{t+1}|\tilde{C}_{i-1}|$ so obtained is no larger than the already existing $|\ell_t|$. In the process we have made $S_i - \ell_{t+1}(v_i)$ smaller by at least 1 than $S_i - \ell_t(v_i)$. It follows that $|\ell_{t+1}| \le |\ell_t|$. If $S_i - \ell_{t+1}(v_i) > |\ell_{t+1}|\tilde{C}_{i-1}|$ we repeat the process and obtain the labelling ℓ_{t+2}, etc. The process is repeated until there is no obvious room for improvement, that is, until we reach the first index s_i for which $S_i - \ell_{s_i}(v_i) \le |\ell_{s_i}|\tilde{C}_{i-1}|$. The labelling ℓ_{s_i} is then the monotone labelling L_i of C_i with pointers $L_i(v_j) = g(i,j)$ referred to previously. The proof that $|L_i| = B(C_i)$ (and hence that $|L_k| = B(C_k) = B(C)$) will be the object of the remainder of this paper.

3. Justification of the Algorithm

In this section we show that the m-labelling L_k of $C = C(m_1,m_2,\cdots,m_k)$ with pointers $g(k,j)$, $1 \le j \le k$, produced by CATBAND satisfies $|L_k| = B(C)$, and that L_k is produced in $O(|C|k)$ time. This will be done in two basic steps. First we show that there exist an m-labelling ℓ of C satisfying $|\ell| = B(C)$. We then use this result to show $|L_k| = B(C)$.

To achieve the first step, we begin by reducing to the case in which our caterpillar $C = C(m_1,m_2,\cdots,m_k)$ satisfies $m_i > 0$ for all i.

Lemma 1. Suppose a caterpillar $C = C(m_1, m_2, \cdots, m_k)$ satisfies $m_i = 0$ for some i, $1 \le i \le k$. Then there exists a m-labelling ℓ of C such that $|\ell| = B(C)$.

Proof: We proceed by induction on k, the case $k = 1$ being obvious. Let $m_i = 0$ for some i, $i \le i \le k$.

Consider the caterpillars $C_1 = C(m_1, m_2, \cdots, m_{i-2}, m_{i-1} + 1)$ and $C_2 = C(m_{i+1} + 1, m_{i+2}, \cdots, m_{k-1}, m_k)$, where $C_1 = K_2$ if $i = 1$ and $C_2 = K_2$ if $i = k$. Let the links of C_1 and C_2 be $v_1, v_2, \cdots, v_{i-1}$, and $w_{i+1}, w_{i+2}, \cdots, w_k$ respectively, where $v_j v_{j+1} \in E(C_1)$, $w_j w_{j+1} \in E(C_2)$ for all j, and the endpoint degree of v_j is m_j for $1 \le j \le i - 2$ and $m_{i-1} + 1$ for $j = i - 1$, while the endpoint degree of w_j is m_j for $i + 2 \le j \le k$ and $m_{i+1} + 1$ for $j = i + 1$.

We proceed to the determination of $B(C)$. Letting $M = \max\{B(C_1), B(C_2)\}$, it follows from $C_1 \subseteq C$ and $C_2 \subseteq C$ that $M \le B(C)$. We can show $B(C) \le M$ as follows. By induction C_1 and C_2 admit m-labellings ℓ_1 and ℓ_2 respectively such that $|\ell_1| = B(C_1)$ and $|\ell_2| = B(C_2)$. We can then use ℓ_1 and ℓ_2 to obtain an m-labelling of C with absolute value M as follows. Observe first that by monotonicity we may assume that $|C_1|$ is the ℓ_1 label of an endpoint of C_1 adjacent to v_{i-1} while 1 is the ℓ_2 label of an endpoint of C_2 adjacent to w_{i+1}. Next define the labelling ℓ_2' of C_2 given by $\ell_2'(x) = \ell_2(x) + |C_1| - 1$ for all $x \in V(C_2)$, so that $|\ell_2'| = |\ell_2|$. Now join C_1 and C_2 to form C by identifying the endpoint of C_1 having ℓ_1 label $|C_1|$ with the endpoint of C_2 having ℓ_2' label $|C_1|$. We then have a natural labelling ℓ of C given by $\ell(x) = \ell_1(x)$ if $x \in C_1$ and $\ell(x) = \ell_2'(x)$ if $x \in C_2$. Note that $|\ell| = \max\{|\ell_1|, |\ell_2'|\} = \max\{|\ell_1|, |\ell_2|\} = M$, so that $B(C) \leq M$. It follows that $B(C) = M$, and since $|\ell| = M$ and ℓ is monotone, the lemma is proved. ∎

In view of the above lemma, we may henceforth suppose that our given caterpillar $C = C(m_1, m_2, \cdots, m_k)$ satisfies $m_i > 0$ for all i.

We now introduce some further notation. Suppose $C = C(m_1, m_2, \cdots, m_s)$ is a caterpillar and ℓ is a monotone labelling of C. A link v_k of C is called <u>ℓ-critical</u> if $|\ell(v_j) - \ell(v_{j+1})| = |\ell(v_j) - \ell(v_{j-1})| = |\ell|$ for all j, $1 \leq j \leq k$. We call ℓ <u>taut</u> if all links of C are ℓ-critical. Note that if C admits a taut monotone labelling ℓ, then ℓ is uniquely determined and $|\ell| = B(C)$. Hence in this case we will sometimes say that C is taut, the labelling ℓ being understood. Finally we call ℓ <u>tight</u> if $\ell(v_{j+1}) = \min\{\ell(v_j) + |\ell|, S_{j+1}\}$ for all j, $0 \leq j \leq s$. Thus ℓ is tight if its pointers are as large as possible subject to its bandwidth, that is, for any monotone labelling ℓ' with $|\ell'| = |\ell|$ we have $\ell'(v_j) \leq \ell(v_j)$ for all j, $1 \leq j \leq s$. Tight and taut m-labellings of caterpillars are illustrated in Figures 1b and 1c.

The following lemma will be useful in the next theorem. A caterpillar C' is called an <u>augmentation</u> of $C = C(m_1, m_2, \cdots, m_s)$ if $C' = C(m_1, m_2, \cdots, m_i + 1, m_{i+1}, \cdots, m_s)$ for some i, $1 \leq i \leq s$.

<u>Lemma 2</u>. Suppose $C = C(m_1, m_2, \cdots, m_k)$ is a taut caterpillar and C' is an augmentation of C. Then we have $B(C') = B(C) + 1$.

<u>Proof</u>: Let $C' = C(m_1, m_2, \cdots, m_{i-1}, m_i + 1, m_{i+1}, \cdots, m_k)$ for some i $1 \leq i \leq k$. Now let ℓ be a taut m-labelling of C and let ℓ' be any labelling of C'. By the pigeon hole principle there exists j, $0 \leq j \leq k - 1$, such that $|\ell'(v_j) - \ell'(v_{j+1})| > |\ell(v_j) - \ell(v_{j+1})| = B(C)$

and hence $B(C') \geq 1 + B(C)$. On the other hand the m-labelling L of C' defined by the pointers $\ell(v_1), \ell(v_2), \cdots, \ell(v_{i-1}), \ell(v_i)+1, \ell(v_{i+1}), \cdots, \ell(v_k)$ satisfies $|L| = |\ell| + 1 = B(C) + 1$. It follows that $B(C') = B(C) + 1$, as required. ∎

We are now ready for our basic result on monotone labellings of caterpillars.

Theorem 1: For any caterpillar $C = C(m_1, m_2, \cdots, m_k)$ there exists a monotone labelling L of C such that $|L| = B(C)$.

Proof: We proceed by induction on $|V(C)|$. Suppose that the theorem holds for all caterpillars having at most $|V(C)| - 1$ points, and assume to the contrary that the theorem fails for C.

First we claim that if e is any endpoint of C, then $B(C - e) = B(C)$. Suppose not, so that $B(C - e) < B(C)$, By induction there exists an m-labelling ℓ of $C - e$ such that $|\ell| = B(C - e)$. Let v_r be the link of C to which e is joined. If neither $r = 1$ and $m_1 = 1$ nor $r = k$ and $m_k = 1$ hold, then let L be the m-labelling of C defined by the pointers $\ell(v_1), \ell(v_2), \cdots, \ell(v_{r-1}), \ell(v_r) + 1, \ell(v_{r+1}) + 1, \cdots, \ell(v_k) + 1$. Suppose then that $r = 1$ and $m_1 = 1$ or $r = k$ and $m_k = 1$, and by symmetry we may take $r = 1$ and $m_1 = 1$. Now in $C - e$, v_1 is an endpoint attached to the left-handmost link, and hence we may adjust the labelling ℓ to obtain a labelling ℓ' of $C - e$ such that $|\ell'| = |\ell|$ and $\ell'(v) = 1$. We then let L be the m-labelling of C defined by the pointers $\ell'(v_1) + 1, \ell'(v_2) + 1, \cdots, \ell'(v_k) + 1$. In either case we get $|L| \leq |\ell| + 1 = B(C - e) + 1 \leq B(C)$, so that $|L| = B(C)$. Combining this with L being monotone we obtain a contradiction. Thus we have $B(C - e) = B(C)$ for all endpoints e of C.

Next observe that for all endpoints e of C no m-labelling of $C - e$ can be taut. For if not, then by Lemma 2 we get $B(C) > B(C - e) = B(C)$, a contradiction.

Now let e be any endpoint of C, and let ℓ be an m-labelling of C - e such that $|\ell| = B(C - e)$. By the above ℓ is not taut, and by properly adjusting its pointers we may assume that ℓ is tight. Let v_j be the link of C to which e is joined, and let v_t be a link of C - e which is not critical.

Suppose first that t = j so that v_j is not ℓ-critical. Thus we have $|\ell(v_t) - \ell(v_{t+1})| < B(C - e)$ or $|\ell(v_t) - \ell(v_{t-1})| < B(C - e)$. If $|\ell(v_t) - \ell(v_{t+1})| < B(C - e)$ then let L be the m-labelling of C with pointers $\ell(v_1),\ell(v_2),\cdots,\ell(v_j),\ell(v_{j+1}) + 1,\ell(v_{j+2}) + 1,\cdots,\ell(v_k) + 1$, while if $|\ell(v_t) - \ell(v_{t+1})| = B(C - e)$ and $|\ell(v_t) - \ell(v_{t-1})| < B(C - e)$ define L on C by the pointers $\ell(v_1),\ell(v_2),\cdots,\ell(v_{t-1}),\ell(v_t) + 1,\ell(v_{t+1}) + 1,\cdots,\ell(v_k) + 1$. In either case we get $|L| \leq |\ell| = B(C - e) = B(C)$, again yielding the contradiction of an m-labelling of C with $|L| = B(C)$. Hence v_j is ℓ-critical.

We may therefore suppose that t > j or t < j. We only consider the case t > j for brevity, as the case t < j is handled by a symmetrical argument. We take t to be minimal with respect to v_t being not critical and t > j. For any m-labellings of C - e the labels of v_i and its adjacent endpoints are $S_{i-1} + 1,S_{i-1} + 2,\cdots,S_{i-1} + m_i + 1$ for $1 \leq i \leq j - 1$, $S_{i-1} + 1,S_{i-1} + 2,\cdots,S_{i-1} + m_i$ for i = j, and $S_{i-1} - 1,S_{i-1},\cdots,S_{i-1} + m_i$ for $j + 1 \leq i \leq k$. Accordingly we let $S_i' = S_i$ for $1 \leq i \leq j - 1$ and $S_i' = S_i - 1$ for $j \leq i \leq k$. Writing $C(m_1',m_2',\cdots,m_k')$ for C - e (where $m_i' = m_i$ for $i \neq j$ and $m_j' = m_j - 1$), it follows that the subcaterpillar $C_i' = C(m_1',m_2',\cdots,m_i')$ of C - e receives the labels 1 through S_i' under any m-labelling of C - e.

Suppose first that there is some r, $j < r \leq t$, for which $\ell(v_r) = S_{r-1}' + 1$. Let r be minimal with respect to this property and to j < r. By minimality of t all points v_m, $j \leq m \leq t - 1$, are ℓ-critical, and in particular all v_m are ℓ-critical for $j \leq m \leq r - 1$.

There are now two possibilities. Either all v_m are ℓ-critical for $1 \le m \le j - 1$, or there is some s, $1 \le s \le j - 1$, such that v_s is not ℓ-critical.

Suppose first that all the v_m are ℓ-critical for $1 \le m \le j - 1$. Combining this with v_m being ℓ-critical for $j \le m \le r - 1$ and $\ell(v_r) = S'_{r-1} + 1$ it follows that C - e contains the taut subcaterpillar $T = C(m_1, m_2, \cdots, m_{j-1}, m_j - 1, m_{j+1}, \cdots, m_{r-1} + 1)$ having bandwidth B(C - e). Here T is the subcaterpillar of C - e induced by the links $v_1, v_2, \cdots, v_r$ and all endpoints of C - e adjacent to any of $v_1, v_2, \cdots, v_{r-1}$. Thus v_r is viewed as an endpoint (of v_{r-1}) in T. Since $m_i > 0$ for all i, there is an endpoint of e' of C adjacent to v_t. By a previous reduction $B(C - e') = B(C - e) = B(C)$, and by induction there exists a monotone labelling ℓ' of C - e' such that $|\ell'| = B(C - e')$. But C - e' contains the augmentation T + e of T obtained by adjoining the endpoint e to the link v_j of T. Since T was taut, it follows that $B(C - e') \ge B(T + e) = B(T) + 1 = B(C - e) + 1 = B(C - e') + 1$, a contradiction.

If there is some s, $1 \le s \le j - 1$, such that v_s is not ℓ-critical in C - e, then we may reduce to the possibility discussed above as follows. We may without loss assume that s is maximal with respect to $s < j$ and v_s not being ℓ-critical. Since v_{s+1} is ℓ-critical we have $\ell(v_{s+1}) - \ell(v_s) = B(C - e)$. Now since v_s is not ℓ-critical we must have $\ell(v_s) - \ell(v_{s-1}) < B(C - e)$. But since ℓ is tight this can only happen if $\ell(v_s) = S'_s$. Since also $\ell(v_r) = S'_{r-1} + 1$, it follows that C - e contains the taut subcaterpillar $Q = C(m_{s+1} + 1, m_{s+2}, \cdots, m_{r-2}, m_{r-1} + 1)$ having bandwidth B(C - e). Here Q is the subcaterpillar of C - e induced by the links $v_s, v_{s+1}, \cdots, v_r$ and all endpoints of C - e adjacent to any of $v_{s+1}, v_{s+2}, \cdots, v_{r-1}$. Thus v_s

is an endpoint of v_{s+1} and v_r is an endpoint v_{r-1} in Q. As above there is some endpoint e' of v_t such that the caterpillar C - e' contains the augmentation Q + e of Q obtained by adjoining the endpoint e to v_j in Q. Combining this with the tautness of Q in C - e and Lemma 2 we obtain the contradiction $B(C - e') \geq B(Q + e) = B(Q) + 1 = B(C - e) + 1 = B(C - e') + 1$.

This completes the discussion of the case in which there is an r, $j < r \leq t$, such that $\ell(v_r) = S'_{r-1} + 1$. We may therefore suppose that $\ell(v_m) > S'_{m-1} + 1$ for all m, $j < m \leq t$. It follows that we may define an m-labelling M of C with pointers $\ell(v_1), \ell(v_2), \cdots, \ell(v_t), \ell(v_{t+1}) + 1, \ell(v_{t+2}) + 1, \cdots, \ell(v_k) + 1$. Clearly $|M(v_m) - M(v_{m+1})| = |\ell(v_m) - \ell(v_{m+1})|$ for $m \neq t$, $0 \leq m \leq k$, and $|M(v_t) - M(v_{t+1})| = |\ell(v_t) - (\ell(v_{t+1}) + 1)| = 1 + |\ell(v_t) - \ell(v_{t+1})| \leq B(C - e) = B(C)$ since v_t is not ℓ-critical. Thus M is an m-labelling of C such that $|M| = B(C)$, a contradiction. The theorem is thus proved. ∎

We are now ready to complete the justification of the algorithm CATBAND. Given the caterpillar $C = C(m_1, m_2, \cdots, m_k)$, let C_i denote the subcaterpillar $C_i = C(m_1, m_2, \cdots, m_i)$ for $1 \leq i \leq k$. The algorithm produces an m-labelling of C_i, denoted ℓ_{s_i} in the discussion above, defined by the pointers $g(i,j)$, $1 \leq j \leq i$, for each i, $1 \leq i \leq k$.

Theorem 2: Let $C = C(m_1, m_2, \cdots, m_k)$ be an arbitrary caterpillar. For any i, $1 \leq i \leq k$, the m-labelling ℓ_{s_i} of C_i defined by the pointers $g(i,j)$, $1 \leq j \leq i$, satisfies $|\ell_{s_i}| = B(C_i)$. In particular we have $B(C) = |\ell_{s_k}|$.

Proof: For simplicity suppose that the construction of ℓ_{s_i} by CATBAND makes use of step 2 but not step 1. If step 1 were operative then certain unimportant complications enter our argument, and hence we will omit this possibility here.

Set $L_i = \ell_{s_i}$. We will prove the stronger assertion that $|L_i| = B(C_i)$ and that L_i is tight for all i, using induction on i. The case i = 1 follows immediately from $B(K_{1,i}) = [\frac{i}{2}]$ and $g(1,1) = \{\frac{i}{2}\}$. Suppose that $|L_j| = B(C_j)$ and that L_j is tight for all $j < i$, where $1 < i \leq k$. Let $B(C_i) = B(C_{i-1}) + x$ for some $x \geq 0$, and as in CATBAND we let $f_j = B(C_j)$ for $1 \leq j \leq i$.

We first recall some notation necessary for the induction. Referring to the discussion following the algorithm and to the use of step 2 in CATBAND we denote by ℓ_0 the m-labelling of C_i with pointers $L_{i-1}(v_i), L_{i-1}(v_2), \cdots, L_{i-1}(v_{i-1}), \min\{L_{i-1}(v_{i-1}) + f_{i-1}, S_i\}$ (or equivalently, pointers $g(i-1,1), g(i-1,2), \cdots, g(i-1,i-1), \min\{g(i-1,i-1) + f_{i-1}, S_i\}$). Referring again to this discussion we let ℓ_x be the m-labelling of C_i with pointers $\ell_x(v_1) = \min\{\ell_0(v_1) + x, S_1\}$, and $\ell_x(v_j) = \min\{\ell_x(v_{j-1}) + f_{i-1} + x, S_j\}$ for $1 < j \leq i$.

First we show that $|\ell_x| = B(C_i)$. By induction $|L_{i-1}| = f_{i-1}$ and L_{i-1} is tight on C_{i-1}. It follows that $\ell_0(v_j) \geq R(v_j)$, $1 \leq j \leq i-1$, for any m-labelling R of C_{i-1} satisfying $|R| \leq f_{i-1}$. Now by theorem 1 there exist an m-labelling M of C_i such that $|M| = B(C_i) = f_{i-1} + x$. Since obviously $|M|C_{i-1}| \leq |M| = f_{i-1} + x$ we get $\ell_x(v_j) \geq M(v_j)$ for all $1 \leq j \leq i$. Hence we get $S_i - \ell_x(v_i) \leq S_i - M(v_i) \leq B(C_i)$. Since the definition of ℓ_x immediately gives $\ell_x(v_j) - \ell_x(v_{j-1}) \leq B(C_i)$ for $1 \leq j \leq i$, it follows that $|\ell_x| \leq B(C_i)$. The opposite inequality $|\ell_x| \geq B(C_i)$ is immediate from the fact that ℓ_x is a labelling of C_i.

Next we show that the labelling $L_i = \ell_{s_i}$ is in fact the labelling ℓ_x, that is, that $\ell_x = \ell_{s_i}$. In our discussion we denoted by ℓ_t the m-labelling of C_i defined by the pointers $g_t(i,j)$, $1 \leq j \leq i$, obtained in the i'th iteration of the algorithm. Hence the pointers of ℓ_x are, in the language of the algorithm, $g_x(i,1), g_x(i,2), \cdots, g_x(i,i)$. Recall that s_i is defined as the least nonnegative integer such that $S_i - g_{s_i}(i,i) \leq f_{i-1} + s_i$. Since $S_i - g_x(i,i) \leq |\ell_x| = B(C_i) = f_{i-1} + x$, we have $x \geq s_i$. Suppose that $x > s_i$. If $\ell_x = \ell_{s_i}$ then there is nothing to prove, so assume $\ell_x \neq \ell_{s_i}$. Then by tightness of L_{i-1} and the definition of ℓ_x there exists j, $1 \leq j \leq i$, such that $\ell_x(v_j) - \ell_x(v_{j-1}) > f_{i-1} + s_i = |\ell_{s_i}|$. Thus we get $|\ell_x| > B(C_i)$, contradicting $|\ell_x| = B(C_i)$, and hence $x = s_i$ so $\ell_x = \ell_{s_i} = L_i$ as required. To complete the induction, we note that L_i is tight since, as observed above, $L_i(v_j) = \ell_x(v_j) \geq M(v_j)$ and M was an arbitrary m-labelling of C_i satisfying $|M| = f_{i-1} + x = |\ell_x|$. We have thus shown that the m-labelling L_i produced by CATBAND satisfies $|L_i| = B(C_i)$ and the theorem is proved. ∎

Next we verify that CATBAND is indeed a polynomial algorithm.

Theorem 3: The algorithm CATBAND determines the bandwidth of the caterpillar $C = C(m_1, m_2, \cdots, m_k)$ in at most $|C|k$ steps.

Proof: Recall that CATBAND determines successively the bandwidths $B(C_i)$ of the subcaterpillars $C_i = C(m_1, m_2, \cdots, m_i)$, $1 \leq i \leq k$, of C.

By theorem 2 we know that the integer $f_i = |\ell_{s_i}|$ is indeed the bandwidth $B(C_i)$ of C_i. Let us then analyze the number of steps required in determining f_i given f_{i-1}. Recall that $f_i = \lambda_{i-1} + s_i$ where $\lambda_{i-1} = g_{i-1} + 1$ or f_i depending on whether step 1 or step 2 respectively is used. Hence we are reduced to determining the time required for finding the integer s_i. By definition s_i is the smallest nonnegative integer for which $S_i - g_{s_i}(i,i) \leq \lambda_{i-1} + s_i$. Hence the determination of s_i requires the successive construction of the pointers $\{g_t(i,1), g_t(i,2), \cdots, g_t(i,i)\}$ (of the m-labelling ℓ_t), for all $t \geq 0$ until we reach the smallest t satisfying $S_i - g_t(i,i) \leq \lambda_{i-1} + t$. Now the construction of the set of pointers $\{g_t(i,j): 1 \leq j \leq i\}$ given the set $\{g_{t-1}(i,j): 1 \leq j \leq i\}$ requires i steps, one step used for finding $g_t(i,j)$ for each j by using the formula $g_t(i,j) = \min\{g_t(i,j-1) + \lambda_{i-1} + t, S_i\}$. The parameter t can take on at most $S_i - S_{i-1} = 1 + m_i$ values. Hence we use at most $i(m_i + 1)$ steps in finding f_i given f_{i-1}.

To finally determine $B(C) = B(C_k)$ we must repeat the inductive step described above for each integer i, $1 \leq i \leq k$. For convenience let $n = |C|$. Hence the total number of steps used is bounded above by

$$\sum_{i=1}^{k} (1 + m_i)i \leq \frac{k(k+1)}{2} + k(n-k) \leq kn, \quad \text{since} \quad \sum_{i=1}^{k} m_i = n - k.$$

It follows that CATBAND requires at most $|C|k$ steps to determine B(C), and the theorem is thus proved. ▓

The following is a simple consequence of Theorem 2.

<u>Corollary 1</u>: For any caterpillar $C = C(m_1, m_2, \cdots, m_k)$ we have $B(C) = \max\{\frac{|C'| - 1}{d(C')}\}$, where the maximum is taken over all subcaterpillars C' of C, and d(C') is the diameter of C'.

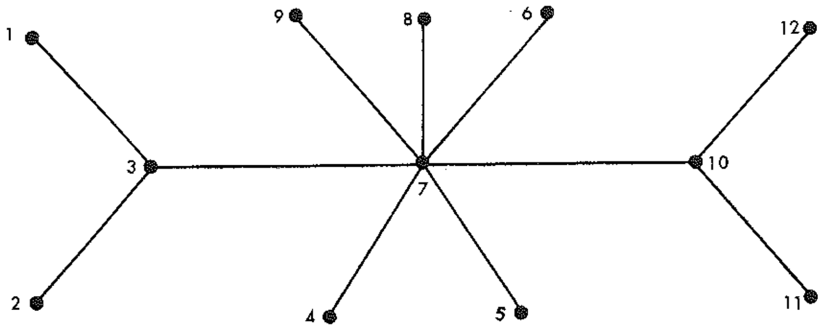
In the next corollary we see that our result on caterpillars can be used in finding the bandwidth of other graphs. Let $C' = C'(m_1, m_2, \cdots, m_k)$ be the graph obtained from $C = C(m_1, m_2, \cdots, m_k)$ by adding all possible edges joining pairs of endpoints of C attached to the same link v_j, for each j, $1 \leq j \leq k$. Thus C' has the same underlying point set as V(C), but the graph induced by the link v_j and its non-link neighbors is K_{1,m_j} in C while it is K_{m_j+1} in C'.

Corollary 2: Let $C' = C'(m_1, m_2, \cdots, m_k)$. Then $B(C') = \max\{B(C), m_j: 1 \leq j \leq k\}$.

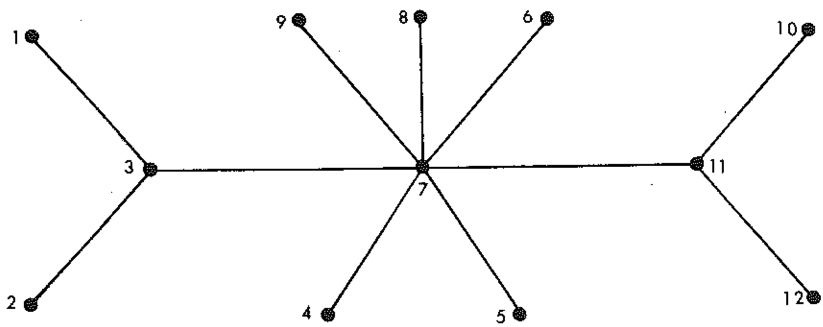
Proof: Set $M = \max\{B(C), m_j: 1 \leq j \leq k\}$.

Since $C \subseteq C'$ we have $B(C') \geq B(C)$. Similarly since $K_{m_j+1} \subseteq C'$ for all j we have $B(C') \geq \max\{m_j: 1 \leq j \leq k\}$, and hence $B(C') \geq M$. On the other hand, the m-labelling $\ell_{s_k} = L_k$ of C produced by CATBAND may be used as a labelling of C' (using the natural correspondence between the points of C and C'). Then obviously the restriction $L_k | K_{m_j+1}$ of L_k to each of the complete subgraphs of C' satisfies $|L_k | K_{m_j+1}| = m_j$ while the restriction $L_k | C$ to C satisfies $|L_k | C| = B(C)$ by theorem 2. It follows that L_k (viewed as a labelling of C') satisfies $|L_k| = M$, and the corollary is proved. ∎

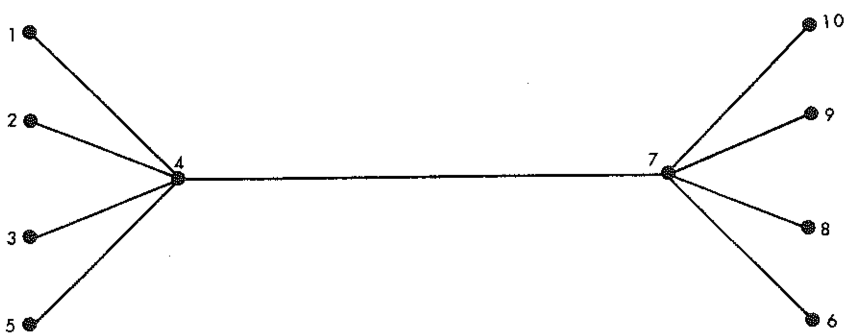
Remark: A O(n log n) algorithm for determining the bandwidth of caterpillar graphs has been obtained independently (and nearly simultaneously) in a paper by Assmann, Kahn, Kleitman, Syslo, and Zak. Their algorithm applies also to caterpillars in which the "hairs" have length 2.



a. An m-labelling of C(2,5,2)



b. A tight m-labelling of C(2,5,2)



c. The taut caterpillar C(4,4).

Figure 1: Monotone, tight, and taut labellings of caterpillars.

References

1. M. Behzad, G. Chartrand, L. Lesniak - Foster - Graphs and Digraphs, Prindle, Weber and Schmidt, Boston (1979).

2. J. A. Bondy, U.S.R. Murty - Graph Theory with Applications, Elsevier North Holland, New York (1976).

3. P. Z. Chinn, J. Chvatalova, A. K. Dewdney, N. E. Gibbs - Graph bandwidth: A survey of theory and applications, submitted.

4. J. Chvatalova - On the bandwidth problem for graphs, Ph.D. dissertation, Department of Mathematics, University of Waterloo, 1980.

5. M. R. Garey, R. L. Graham, D. S. Johnson, D. E. Knuth - Complexity results for bandwidth minimization, SIAM Journal for Applied Mathematics, 34 (1978) 477-495.

6. F. Harary - Graph Theory, Addison-Wesley, Reading (1969).

7. F. Harary - Problem 16, Theory of Graphs and its Applications, p. 161 (M. Fiedler, editor), Czechoslovak Academy of Sciences, Prague (1964).

8. C. H. Papadimitriou - The NP-completeness of the bandwidth minimization problem, Computing 16 (1976) 263-270.